

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.⁷
C12Q 1/68
G01N 33/53



[12] 发 明 专 利 说 明 书

[21] ZL 专利号 96198630.1

[43] 授权公告日 2003 年 8 月 13 日

[11] 授权公告号 CN 1117876C

[22] 申请日 1996.11.27 [21] 申请号 96198630.1

[30] 优先权

[32] 1995.11.30 [33] US [31] 60/007,722

[86] 国际申请 PCT/US96/19009 1996.11.27

[87] 国际公布 WO97/20075 英 1997.6.5

[85] 进入国家阶段日期 1998.5.28

[71] 专利权人 研究发展基金会

地址 美国内华达州

[72] 发明人 M·R·蒙特米尼 B·皮尔斯

[56] 参考文献

MOLECULAR ENDOCRINOLOGY VOL. 8, NO. 12 1994
-01 -01 PEERS ET AL : INSULIN EXPRESSION IN
PANCREATIC ISLET CELLS RELIES ON COOPERATIVE
INTER

审查员 常 矛

[74] 专利代理机构 上海专利商标事务所

代理人 陈文青

权利要求书 2 页 说明书 29 页 附图 14 页

[54] 发明名称 刺激促生长素抑制素和胰岛素产生的化合物的筛选方法

[57] 摘要

本发明提供一种设计模拟 Pbx 使 STF-1 结合稳定的化合物的方法。用熟知的 DNA 结合测定法,该领域普通技术人员可以筛选出化合物来测定有效促进 STF-1 与 DNA 结合的药物。通过这种方法,可以发现新的用于刺激促生长素抑制素和胰岛素产生的化合物。

I S S N 1 0 0 8 - 4 2 7 4

1. 一种 DNA 结合的测定方法，以测定能有效促进 STF-1 与 STF-1 接合部位相结合的化合物，其方法包括步骤：

5 将末端标记的、有 STF-1 结合位点的双链 DNA 与 STF-1 混合，作为第一容器中的对照；

将末端标记的、有 STF-1 结合位点的双链 DNA 与 STF-1 和待测试化合物样品混合，作为第二容器中的测试样品；

培育所述第一和第二容器；

10 将所述对照和所述测试样品上样于电泳凝胶；

对所述电泳凝胶施加一定电流以使所述对照和所述测试样品在所述凝胶中迁移；

检测所述对照和所述测试样品；

15 比较所述对照和所述测试样品的迁移，其中如果所述测试样品的迁移比所述对照慢，则所述测试化合物是有效地促进了 STF-1 与所述 STF-1 结合位点的结合，其中所述 STF-1 结合位点含有在与蛋白辅因子结合时能与 STF-1 特异性结合的 DNA 序列。

2. 根据权利要求 1 所述的测定方法，其中制得的测试样品含有末端标记的有所述 STF-1 结合位点的双链 DNA、STF-1、测试化合物、和过量的有与所述标记过的双链 DNA 相同序列的未标记过的双链 DNA，其中如果所述测试样品的检测导致标记的 DNA/STF-1/测试化合物的复合体减少，则所述 STF-1/测试化合物的复合体的结合对于有所述 STF-1 结合位点的双链 DNA 来说是特异性的。

3. 根据权利要求 1 所述的测定方法，其中所述末端标记过的双链 DNA 中的一条链有 SEQ ID No.1 的序列，另一条链有 SEQ ID No.2 的序列。

25 4. 一种 DNA 结合的测定方法，以测定能有效促进 STF-1 与 STF-1 接合位点相结合的化合物，其方法包括步骤：

将组成型表达 STF-1 的第一表达质粒和在 STF-1 结合位点调制下表达报道基因的第二表达质粒转染入细胞系中；

30 将第三质粒转染入所述转染过的细胞系中，所述第三表达质粒可表达测定化合物；和

测定所述报道基因的转录量，其中如果所述转录发生，则所述测试化合物可有效促进 STF-1 与所述 STF-1 结合位点的结合，其中所述 STF-1 结合位点含有在

与蛋白辅因子结合时能与 STF-1 特异性结合的 DNA 序列。

5. 根据权利要求 4 所述的测定方法, 其中所述第二表达质粒含有一个拷贝的促生长素抑制素 TSE II 位点作为所述 STF-1 结合位点。

6. 根据权利要求 4 所述的测定方法, 其中所述合适的细胞系是 GC 细胞。

5

刺激促生长素抑制素和胰岛素产生的化合物的筛选方法

5

发明背景

发明领域

本发明涉及生物化学内分泌学、分子生物学和蛋白质化学。更具体地，本发明涉及测定刺激促生长素抑制素和胰岛素产生的化合物的方法。

相关技术描述

10

葡萄糖的体内稳态需要多种神经内分泌系统的协同作用。然而，胰岛被认为是哺乳动物体内主要的“葡萄糖感受器”。胰岛包含四种细胞，它们被称为前胰岛素原(primansulin)、胰高血糖素、促生长素抑制素或胰多肽。其中产生胰岛素的 β 细胞占主要地位。血清中葡萄糖的增加刺激了胰岛素的分泌和产生，而该步骤可使某些组织随后吸收葡萄糖。因此， β 细胞的功能异常或破坏将导致血清葡萄糖含量的升高，最终导致糖尿病。

15

遗传连锁分析表明遗传因子对获得糖尿病的易感性有很大影响。例如，至少18个基因座与胰岛素依赖型糖尿病(IDDM)有一定的关联。一个称为IDDM2的疾病感受型基因座包括人的胰岛素基因，并与胰岛素启动子功能的改变的转录调节有关。因此，调节胰岛素基因表达步骤的破坏将部分导致糖尿病原性疾病。与这种假设一致的是， β 细胞功能的损失是糖尿病的一个非常普遍的特征。

20

非胰岛素依赖型糖尿病(NIDDM)被认为是外界影响和复杂的遗传影响两者共同的结果。令人感兴趣的是，胰岛素基因座本身的等位基因变异与糖尿病有关。这些变异表现出含有正常的胰岛素基因，但是表现出改变了与转录调节有关的性质。

25

调查表明有多达2千万的美国人患有II型糖尿病。疾病的发展需要环境因素和某些还未被充分鉴定的可能使II型糖尿病有外周胰岛素抗性的糖尿病易感性基因(diabetes susceptibility gene)，在II型糖尿病中组织就不能响应胰岛素信号来适当地利用葡萄糖。或者，在这些患者中，遗传因子可能是造成产生胰岛素的胰 β 细胞的葡萄糖易感性减少的原因。这两种生理状态的最终结果是显著的高血糖，而高血糖是糖尿病的主要特征。

30

胰岛素基因的转录调控是通过与细胞特异性的、对葡萄糖敏感的信号分子相互作用的短区域侧翼DNA来实现的。尽管已经普遍认为含有碱性螺旋-环-螺旋

(bHLH)和同源异型域的因子是控制 β 细胞特异性表达胰岛素的转录机理的关键成分,但是该调节组织的确切特征仍然知之甚少。对胰岛有特异性的碱性螺旋-环-螺旋复合体与近侧的分别称为 Nir、IEB1 或 ICE 的 E-框(E-box)相互作用;该元件存在于鼠的胰岛素 I 基因中,但在鼠胰岛素 II 和人胰岛素基因中只出现一次。

5 在产胰岛素的细胞系中的瞬时测定(transient assay)表明,与 E 框结合的因子可与对 β 细胞有特异性的蛋白质发生协同作用,该蛋白质与邻近的称作 FLAT 的富含 AT 的序列结合,FLAT 则具有与同源异型域识别序列的特征。与 FLAT 结合的因子包括 Isl-1、lmx-1、cdx-3 和 STF-1。此外,其中的后者与称作 P 元件的进化上保守而富含 AT 的序列有首要的结合活性。Isl-1 与 FLAT 元件的结合很弱,而且用产胰岛素的细胞抽提物检测时,似乎不存在于结合 FLAT 的复合体中。10 现有的数据支持了 Isl-1 在神经生长中所起的更重要的作用。同源异型域因子 lmx-3 和 cdx-3 在异源细胞中有令人感兴趣的胰岛素启动子功能的反式激活性质,但是它们在 β -细胞内的细胞分布和与 FLAT 结合的活性仍是不清楚的。另外,直接表明这些因子在 β 细胞系中的功能的数据也很少。

15 在有与胰岛素启动子结合活性的因子组中,STF-1 可能是最有前途的胰岛素启动子功能的真正调控器。在小鼠中,于胚胎期 8.5 天在形成胰脏的原基细胞核中首先检测到有 STF-1,略早于该区域内最早测到有胰岛素表达之前。随后通过内分泌胰的生长,STF-1 和胰岛素大量共表达。另外,在产胰岛素细胞系的抽提物中,STF-1 是当胰岛素启动子中同时有 FLAT 和 P 元件时有内源 DNA 结合活性的成分。STF-1 也与与 E-框结合的因子 Pan-1(估计可从与 FLAT 结合的因子中20 获得)有强的协同作用。然而,DNA 结合测定法表明,来自 β 细胞抽提物的其它未知因子也对检测到的 FLAT 结合活性有较大的贡献。FLAT 介导的胰岛素启动活性是否需要全部这些检测到的种类、或仅仅需要一个亚型(subset)仍是不清楚的。

25 成熟的胰脏包括内分泌和外分泌两种成分,它们被认为是从生长的肠的普通前体细胞中产生的。该前体细胞表达出称作 STF-1(14,20)(或 IPF-1(19),IDX(17)或 X1Hbox 8(22))的同源异型域蛋白,其重要性在“剔除”研究中有所描述,“剔除”研究是指寻靶破坏 STF-1 基因将导致胰的先天性缺失。尽管 STF-1 最初是由生长的胰的内分泌细胞和外分泌细胞表达的,但是 STF-1 的产生逐渐局限于产生胰岛素和促生长素抑制素的胰岛细胞(7)。在这些细胞中,STF-1 的作用对于维持促生长素抑制素和胰岛素基因的高水平表达显得很重要(14,17,19,20,22)。

30 STF-1 可识别两种确切的胰岛素启动子上的称作 FLAT 和 P 的胰岛特异性元

件。当与这些位点结合时，STF-1与E47协同刺激胰岛素的转录，E47是可识别两种称作Far和Nir的E基因框元件的螺旋-环-螺旋蛋白质。同样的，STF-1通过两个胰岛特异性元件(称为TSEI和TSEII)来调控胰岛细胞中促生长素抑制素的表达(14,17)。

- 5 除了对某些基因有转录刺激作用，发现同源异型域蛋白质如STF-1在确定细胞或节段身份(segmental identity)的研究中有重要作用。与它们在体内的特异性和显著的活性相反，许多同源异型域蛋白质在体外表现出较低的与重叠DNA结合的特异性。然而，目前的研究暗示，某些蛋白质辅因子可作为体内同源异型域DNA结合特异性的决定簇(8, 13)。例如在果蝇属中，外小齿(extradenticle)(exd)
- 10 可调节同源异型蛋白质的活性而不改变它们的表达方式(21,24)。另外，外小齿表现出通过增强某些同源异型域蛋白质的DNA结合特异性来促进靶基因的选择(2,25)。实际上，外小齿在脊椎动物中是高度保守的，其与人原癌基因Pbx1有71%的延伸序列相同(23)。

- 15 已有技术缺少有效的增强同源异型域蛋白如STF-1的结合的方法。本发明满足了该领域中这一长期存在的需要。

发明概述

- 许多同源异型域蛋白表现出可通过刺激特异靶基因的转录来调节细胞生长。与它们在体内的显著活性相比，许多同源异型域蛋白在体外与蛋白质靶部位的结合是无特异性的，这表明有识别靶部位选择性的辅因子参与。一种辅因子称为外小齿(exd)，它可通过与某些同源异型域蛋白协同结合到靶调节元件上来影响果蝇属中的节段形态发生(segmental morphogenesis)。本发明证明了STF-1(一种哺乳动物中胰生长所需要的孤独同源异型域蛋白)通过与Pbx(哺乳动物中外小齿的同源物)协同结合到DNA上来刺激胰岛素基因促生长素抑制素的转录。与Pbx
- 20 的协同结合需要五肽基序(pentapeptide motif)(FPWMK)，它在同源异型域蛋白的亚型中是高度保守的。FPMWK基序不足以使其它同源异型域蛋白具有Pbx协同性，然而，STF-1同源异型域的N末端臂也是必须的。当与Pbx的协同结合发生在仅一个潜在STF-1靶部位亚型上时，本发明表现出Pbx可通过和STF-1形成异二聚复合体来选择性识别胰生长中的靶基因。实际上，本发明证明了促生长素抑制素启动子的TSEII元件可识别胰岛细胞中的由STF-1和Pbx组成的异二聚复合体。因此，胰中STF-1作用的特异性部分受潜在靶启动子部位识别STF-1 Pbx异二聚复合体的能力的支配。
- 30

因此, 本发明的一个方面是提供了一种 DNA 结合测定法来测定对于促进 STF-1 与 STF-1 接合部位的结合有效的化合物, 其步骤包括: 使有末端标记的、有 STF-1 结合位点的双链 DNA 与 STF-1 一起混合, 作为第一容器中的对照; 使末端标记的、有 STF-1 结合位点的双链 DNA 与 STF-1 和测定化合物一起混合, 5 作为第二容器中的样品; 培育所述第一和第二容器; 将所述对照和所述样品上样于电泳凝胶; 对所述电泳凝胶施加一定电流以使所述对照和所述样品在所述凝胶中迁移; 检测所述对照和所述样品; 比较所述对照和所述样品的迁移, 其中如果所述样品的迁移比所述对照慢, 则所述测定的化合物是有效地促进了 STF-1 与所述 STF-1 结合位点的结合。

10 本发明的另一方面是提供了一种 DNA 结合测定法来测定可有效促进 STF-1 与 STF-1 接合部位的结合的化合物, 步骤包括: 将组成型表达 STF-1 的第一表达质粒和在 STF-1 结合位点控制下表达受体基因的第二表达质粒转染入合适的细胞系中; 将第三质粒转染入所述转染合适细胞系中, 所述第三表达质粒可表达测定 (trest) 化合物; 测定所述受体基因的转录量, 其中如果所述转录发生, 则所述测定 15 化合物可有效促进 STF-1 与所述 STF-1 结合位点的结合。

本发明的其它方面、特征和优点可从下面公开给出的本发明的较佳实施例的描述中明显获得。

附图简述

20 为使上述的本发明的特征、优点和目的被详细了解, 本发明以上概述的详细描述可参照一些结合附图来描述的实施例。这些附图是说明书的一部分。然而, 应当注意的是, 附图描述了本发明的较佳实施例, 因此它不能被认为是限制了它们的范围。

图 1 表示胰的同源异型框因子 STF-1 与遍在核因子一起协同结合到促生长素抑制素启动子中的细胞特异性调节元件上。图 1A 显示了胰岛 Tu-6 细胞的粗核抽提物用促生长素抑制素 TSEII 从 -303 延伸至 -281 的双链寡核苷酸进行的凝胶移位分析。其中指出了 C1、C2 和 C3 是蛋白质 DNA 复合体。渐增的条带表明核抽提物的量的增加。STF, 只有重组 STF-1 蛋白质; Tu-6 Ne 和 Hela NE 分别指 Tu-6 25 和 Hela 细胞的核抽提物。PI 指预免疫抗血清。STF-1 Ab 指抗 STF-1 蛋白质的 C 末端部分产生的抗血清。STF-1 Ab+ Oct 寡核苷酸指 STF-1 抗血清加上 SV40 的 Oct1 结合位点 (Sph1 基序)。Oct. Ab 指抗 Oct.1 蛋白质产生的单克隆 Oct 1 抗血清 (W. Herr 赠与)。图 1B 显示了用如下所述的促生长素抑制素 TSEII 和 Tu6 核抽 30

提物制得的复合体(预先与未标记的 Oct1 寡核苷酸培育,以只对 C1 和 C2 复合体分析)的解离速率(off-rate)分析。复合体 C1 和 C2 如图所示。每条泳道均显示了在加入 1000 倍过量的未标记 TSEII 竞争剂 DNA 后的时间(分钟表示)。PI 指预免疫的血清; STF-1 Ab 指加入凝胶移动测定中的 STF-1 抗血清。图 1C 显示了重组 STF-1 蛋白质的凝胶移位分析,其中用促生长素抑制素 TSE II 寡核苷酸作为探针。STF-1 结合活性在没有(-)或有(+)Jurkat 细胞的异源核抽提物(用 Western 印迹法检测,不含可检测水平量的 STF-1 蛋白质)下测定。渐增条带表明重组 STF-1 蛋白质的量的增加。复合体 C1 和 C2 如图所述。图 1D 显示了含有重组 STF-1 加上 Jurkat 核抽提物的重建抽提物中复合体 C1 和 C2 的解离速率。每条泳道均显示了在加入未标记 TSE II 竞争剂(1000 倍过量)后的时间点(分钟表示)。

图 2 表示遍在同源异型框蛋白质 Pbx 与 STF-1 在促生长素抑制素 TSEII 元件上形成异二聚复合体。图 2A 显示了用 TSE II 寡核苷酸作为探针对 Tu6 NE 的凝胶移位分析。复合体 C1、C2 和 C3 如图所示。PBX Ab 指反应中包含的 Pbx 抗血清。PI 指预免疫血清。STF-1 Ab 指 STF-1 抗血清。图 2B 显示了体外翻译的 Pbx 蛋白质(PBX)对于 STF-1 结合活性的影响。如每条泳道中所示,凝胶移位分析采用 TSE II 寡核苷酸加上 STF-1 和/或 Pbx。TNT 指对照的裂解液。图 2C 显示了 PBX 使 STF-1 与 TSEII 复合体的结合稳定。STF-1 和 STF-1 Pbx 复合体的解离速率分析用促生长素抑制素 TSE II 寡核苷酸作为探针。加入过量未标记 TSE II 寡核苷酸后的时间(分钟表示)如图所示。如图所述,复合体 C1 和 C2 分别与 STF-1 单体和 STF-1/PBX 异二聚体对应。图 2D 显示在促生长素抑制素和胰岛素启动子序列上 STF-1/Pbx 异二聚体的形成的分析结果。TSEII 指野生型 TSE II 寡核苷酸。图 3D 和 3E 显示了 M1、M2 和 M3 以及突变型 TSEII 寡核苷酸(TAAT 基序发生突变)。P 和 FLAT 是可高度亲和识别 STF-1 的胰岛素 I 元件。TSEI 是与 STF-1 结合的促生长素抑制素启动子元件。PBX 指加入结合反应中的体外翻译 PBX 蛋白; STF 指全长重组 STF-1 蛋白质。图 2E 显示了图 3B 的凝胶移位分析中采用的野生型和突变型寡核苷酸。连括号表示诱变寻靶的共有序列 TAAT 基序。图 2F 显示了 STF-1 和 Pbx 协同作用在启动子位点的亚型上。每条线条下指出了用 STF-1 和 E2A-Pbx 效应质粒进行的 GC 细胞瞬时转染测定法。线条框图显示了从受体构建物获得的萤光素酶活性,受体构建物包括含有单个的最小生长激素(GH)的启动子(Luc),两个拷贝的 GH 启动子上游的促生长素抑制素 TSEII 元件(TSEII-Luc)、或两个拷贝的胰岛素 P 元件(P-Luc)。使活性标准化以共转染入 RSV-CAT 对照质粒中。

图 3 表示 STF-1 中保守的五肽基序对于与 Pbx 的协同结合是关键。图 3A 显示了野生型和截短重组的 STF-1 多肽形成的 STF-1 单体(复合体 C1)和 STF-1/Pbx 异二聚体(复合体 C2)的分析结果。每条泳道中指出了突变的 STF-1 多肽中缺失端点。例如, D1-70 指出 STF-1 多肽缺少残基 1-70。Hox140-215 指 STF-1 的同源异型域多肽。图 3B 显示上述凝胶移位测定中所用的构建物的示意图。H.D. 指 STF-1 同源异型域(氨基酸残基 140-215)。全长 STF-1 蛋白质从氨基酸残基 1 延伸至 284。图 3C 显示保守的五肽基序的诱变破坏了与有所示氨基酸数目的野生型和突变型 STF-1 蛋白质的 Pbx 序列协同性。图 3D 显示了单独的以及体外翻译 PBX 或 E. Jurkat 核抽提物组合使用的野生型(WT)和突变型(MUT)STF 蛋白的凝胶移位分析。

图 4 表示 STF-1 内的五肽基序(其对于 STF-1 与 Pbx 在促生长素抑制素 TSEII 位点上形成异二聚体是必须的)在多个同源异型框蛋白中是保守的。不同的同源异型框蛋白根据种类来分类并根据名字排列。对于每种蛋白质,与 STF-1 中基序对应的序列以单字母氨基酸码表示,右侧显示了 N 端至同源异型域的距离。

图 5 表示与 PBX 相互作用的保守基序(Pim)对于促进 PBX 和同源异型框蛋白质间的协同结合是必须的,但量不足。图 5A 显示了下面凝胶移位测定中所用的重组 GST-融合蛋白质的示意图。Isl1 H.D.指 Isl1 因子 lim 域的同源异型域。PBX 相互作用基序-Isl1 H.D.指含有融合入 Isl1 同源异型域的 STF-1 PBX 相互作用基序区(氨基酸残基 110-138)的 STF-1-Isl1 融合蛋白。cdx3 H.D.指胰同源异型框蛋白 cdx3 的同源异型域(氨基酸残基 143-253)。PBX 相互作用基序-cdx3 指含有与 cdx3 同源异型域(氨基酸 176-253)融合的 STF-1 PBX 相互作用基序区(氨基酸 110-138)的融合蛋白。PBX 相互作用基序-STF-1 指含有与 STF-1 同源异型域(氨基酸残基 140-215)融合的 PBX 相互作用基序区(110-138)的 STF-1 多肽。STF-1 H.D. 指单独 STF-1 同源异型域而没有 PBX 相互作用基序区。图 5B 显示了用促生长素抑制素 TSEII 位点作为 ^{32}P 标记的探针进行的凝胶移位测定。 -和+表示没有或含有用 PBX RNA 控制的网织红细胞裂解液。每条泳道显示了用同源异型域融合蛋白测得的 PBX 协同性。单体和异二聚复合体如标记所示。

图 6 表示 STF-1 同源异型域的 N 末端臂是与 Pbx 协同性所必需的。其显示了 PBX 相互作用基序-cdx3 融合构建物的凝胶移位分析结果,该构建物含有 STF-1 PBX 相互作用基序(氨基酸残基 110-138)以及 STF-1 同源异型域的各个被 cdx3 的同源异型域取代的区域。图 6A 显示了 STF-1/cdx3 融合构建物的结构和活性。 cdx3 的序列为阴影部分,而 STF-1 的序列为白色部分。并指出了 N 末端臂(N)和

Helices 1,2 和 3(H1,H2,H3) 的相对位置。凝胶移位测定中与 Pbx 的协同结合在右侧表示成(+,-)。图 6B 显示了 STF-1 和 cdx3(显示在下面)同源异型域的氨基酸排列, 右侧显示了 C 端残基的氨基酸数。破折号表示 STF-1 和 cdx3 间的氨基酸身份。箭头所指地方是用于融合构建物的氨基酸端点(I, II, III)。图 6C 显示了重组的 STF-1/cdx3 融合蛋白的凝胶移位测定结果。构建物数字指上面示意图中所述的构建物。如图所示, 体外翻译的 Pbx(PBX)或未控制的网织红细胞裂解液(TNT)存在于结合反应中。C1、C2 复合体分别与 STF-1 单体和 Pbx/STF-1 异二聚体对应。

10 发明详述

人的基因组含有四簇称为 Hox 基因的同源异型选择者基因, 它们是胚胎发育中轴体模式形成(axial body pattern formation)的关键决定簇(Krumlauf, 1994,Cell 78:191-201)。四个簇各含有多达 13 个基因, 而一个簇中的一个给定基因通常与其它三个家族中的一个有高度的同源性。这种对应的基因称为共生同源基因(paralog); 因此 HoxA1, HoxB1, HoxC1 和 HoxD1 均是紧密对应的共生同源基因, 而每个在不同的簇的不同的染色体上。例如, HoxB 复合体位于染色体 17 的长臂上, 且 HoxB1 至 HoxB9 已被鉴定。

胰岛素基因的葡萄糖依赖性调控随着胰岛素分泌中葡萄糖介导的增加而产生。这可以部分地源于胞间钙的增加。另外, 响应葡萄糖的胰岛素启动子功能可至少部分地通过调节 FLAT 结合蛋白的活性来产生。HoxB13 与胰岛素启动子的有重要功能的 FLAT 元件高亲和性地结合。此外, HoxB13 和胰岛素 ICE/Nir 元件-结合因子 Pan-1 在组合加入时大大地激活了胰岛素启动子。这与 FLAT 和 Nir 元件协同作用于产生胰岛素的细胞的结果是一致的。这些收集数据表明, 钙依赖性信号途径可调节 HoxB13 的功能。

25 根据本发明, 可采用常规的分子生物学、微生物学和重组 DNA 技术领域中的已有技术。这些技术在文献中有充分描述。例如见 Maniatis, Fritsh& Sambrook, “分子克隆实验手册(Molecular Cloning: A Laboratory Manual(1982); “DNA 克隆实践方法(DNA Cloning: A Practical approach),” 第 I 和 II 卷(D.N.Glover ed. 1985); “寡核苷酸合成(Oligonucleotide Synthesis)” (M.J. Gait ed. 1984); “核酸杂交(Nucleic Acid Hybridization)” [B.D. Hames & S.J. Higgins eds. (1985)]; “转录和翻译(Transcription and Translation)” [B.D. Hames & S.J. Higgins eds. (1984)]; “动物细胞培养(Animal Cell Culture)” [R.I. Freshney, ed. (1986)]; “固定化细胞和酶

(Immobilized Cells And Enzymes)” [IRL Press, (1986)]; B. Perbal, “分子克隆指导(A Practical Guide To Molecular Cloning)” (1984).

因此, 如果在这里出现, 下述术语应有下列定义。

- 5 这里所述的氨基酸最好是“L”异构型的。然而, “D”异构型的残基可代替任何L-氨基酸残基, 只要多肽保留有与免疫球蛋白结合的所需功能性质。NH₂指多肽的氨基末端存在的游离氨基基团。COOH指存在于多肽的羧基末端的游离的羧基基团。为与标准的多肽术语(J Biol. Chem., 243:3552-59(1969))保持一致, 下面的对应表显示了氨基酸残基的缩写:

对应表

符号		氨基酸
1 个字母	3 个字母	
Y	Tyr	酪氨酸
G	Gly	甘氨酸
F	Phe	苯丙氨酸
M	Met	甲硫氨酸
A	Ala	丙氨酸
S	Ser	丝氨酸
I	Ile	异亮氨酸
L	Leu	亮氨酸
T	Thr	苏氨酸
V	Val	缬氨酸
P	Pro	脯氨酸
K	Lys	赖氨酸
H	His	组氨酸
Q	Gln	谷氨酰胺
E	Glu	谷氨酸
W	Trp	色氨酸
R	Arg	精氨酸
D	Asp	天冬氨酸
N	Asn	天冬酰胺
C	Cys	半胱氨酸

- 10 应当注意, 所有的氨基酸残基序列在这里以式子表示, 其从左到右是按常规

的氨基端到羧基端的方向排列的。而且，应当注意，氨基酸残基序列的开始或结尾处的破折号表示与另一有一个或多个氨基酸残基的序列相连的肽键。上述表格与这里出现的三个字母和一个字母的记号对应。

“复制子”是任何在体内有自动复制 DNA 功能的遗传元件(如质粒、染色体、病毒)，即在其控制下能够进行复制。

“载体”是一种复制子，如质粒、噬菌体或粘粒，另一个 DNA 片段可与其连接从而来复制所连着的片段。

“DNA 分子”指单链或双链螺旋型的脱氧核苷(腺嘌呤、鸟嘌呤、胸腺嘧啶或胞嘧啶)的聚合体形式。该术语仅仅指分子的一级和二级结构，不局限于任何特定的三级结构。因此，该术语包括所发现的双链 DNA、inter alia、线性 DNA 分子(如限制片段)、病毒、质粒和染色体。在讨论这里的结构时，根据常规只给出了沿 DNA 非转录链(即有与 mRNA 相同序列的链)的 5'到 3'的序列。

“复制起点”指参与 DNA 合成的 DNA 序列。

DNA “编码序列”指在合适的调节序列控制下在体内转录并翻译成多肽的双链 DNA。编码序列的界限是 5'端(相当蛋白质的氨基端)的起始密码子和 3'端(相当蛋白质的羧基端)的翻译终止密码子。编码序列包括(但不局限于)原核细胞序列、来自真核细胞 mRNA 的 cDNA 甚至是合成的 DNA 序列。聚腺苷酸信号和转录终止序列通常位于编码序列的 3'端。

控制转录和翻译的序列是 DNA 调控序列，如启动子、增强子、聚腺苷酸信号、终止子等，它们可提供宿主细胞中的编码序列的表达。

“启动子序列”是可以与细胞中 RNA 聚合酶结合并引发下游(3'方向)编码序列转录的 DNA 调控区。为详细说明本发明，启动子序列的 3'端与转录起始部位连接，其上游(5'方向)延伸包括在上述背景下引发可检测水平的转录所需的最小数量的碱基或元件。在启动子序列中会发现转录起始位点(常规情况下用核酸酶 S1 作图来确定)，还有与 RNA 聚合酶结合的蛋白质结合区域(共有序列)。真核生物启动子通常(但不总是)含有“TATA”框和“CAT”框。原核生物启动子除-10 和-35 共有序列外还含有 SD 序列。

“表达控制序列”是控制并调节另一 DNA 序列转录和翻译的 DNA 序列。当 RNA 聚合酶将编码序列转录成 mRNA 时，编码序列是处于转录和翻译控制序列的“控制”下，然后 mRNA 被翻译成由编码序列编码的蛋白质。

“信号序列”可以包括编码序列之前的部分。该序列编码信号肽，即多肽的 N 末端，其与宿主细胞联系并使多肽到达细胞表面或分泌入基质，该信号肽在蛋

白质离开细胞前被宿主细胞切下。信号序列可与多种原核和真核生物天然蛋白质关联。

术语“寡核苷酸”在这里指本发明的探针，其被定义为包括两个或多个核糖核苷酸的分子，最好超过三个。其确切大小与许多因素有关，而这些因素反过来与寡核苷酸的最终功能和用途有关。

术语“引物”在这里指一种寡核苷酸，无论其是天然存在于纯化的限制性消化物中或是合成的，当在诱导引物延伸产物的合成(与核酸链互补)的条件下，即在有核苷酸、诱导剂如 DNA 聚合酶和合适的温度 pH 下，它可作为合成的引发点。引物可以是单链或双链的，且必须有足够的长度，以在诱导剂存在时引发合成所需的延伸产物。引物的确切长度依赖于许多因素，其包括温度、引物来源及其使用方法。例如，对于诊断应用，根据靶序列的复杂性，寡核苷酸引物通常含有 15-25 或更多的核苷酸，尽管它可含有较少的核苷酸。

这里的引物的选择应“基本上”与特定靶 DNA 序列的不同链互补。这意味着引物必须足够互补来与它们各自的链杂交。因此，引物序列不需要反映模板的确切序列。例如，非互补的核苷酸片段可与引物的 5'端连接，而引物序列的其余部分与链互补。或者，非互补的碱基或更长的序列可散布在引物中，只要引物序列与序列有足够的互补性或其杂交，从而制成延伸产物的合成模板。

如这里所用的，术语“限制性内切核酸酶”和“限制性酶”均指在特异性核苷酸序列处或附近切开双链 DNA 的细菌酶。

当这种 DNA 被引入细胞内，细胞就被外源或异源 DNA “转化”。转化的 DNA 可以或不整合(共价连接)入细胞的基因组。例如在原核生物、酵母和哺乳动物细胞中，转化 DNA 可以维持在如质粒那样的附加型元件中。对于真核细胞，稳定转化的细胞中转化的 DNA 已整合入染色体中，这样就能通过染色体的复制遗传给子代细胞。这一稳定性是由真核细胞建立的由一群含有转化 DNA 的子代细胞组成的细胞系或克隆系的能力来证明。“克隆系”是一群从单个细胞或单一祖细胞(monoclonal ancestor)通过有丝分裂获得。“细胞系”是可以在体外稳定生长许多代的原始细胞的克隆系。

当 DNA 序列确定长度内至少约 75 % (较佳至少约 80 %，最佳至少约 90 或 95 %)的核苷酸相同时，两个 DNA 序列是“基本上同源的”。基本上同源的序列可通过用标准软件比较序列来鉴定，其中序列从序列数据库或 Southern 杂交实验在例如该特定系统确定的严格条件下获得。确定的合适杂交条件是该领域已知的。例如见 Maniatis 等人，同上；DNA Cloning, Vols. I&II, 同上；Nucleic Acid

Hybridization, 同上.

DNA 构建物的“异源”区是在较大 DNA 分子中可鉴别的 DNA 片段, 其在自然界中没有发现与较大分子连接。因此, 当异源区编码哺乳动物基因时, 基因通常被 DNA 侧接(f flank), 而该 DNA 不与源有机物基因组中的哺乳动物基因组 DNA 侧接。在另一个例子中, 编码序列是一个构建物, 其中的编码序列在自然界中没有发现(如含有内含子的基因组编码序列 cDNA, 或含有与天然基因不同的密码子的合成序列)。等位变异或天然发生的突变不会产生这里定义的 DNA 异源区。

这些研究中最常用的标记是放射性元素、酶、暴露在紫外光下时产生荧光的化学物质等。许多荧光性物质是已知的, 并可用作标记。它们包括例如荧光素、若丹明、金胺、得克萨斯红(Texas Red)、AMCA 蓝和 Lucifer 黄。特别的检测物是在山羊中制备的与荧光素通过异硫氰酸酯共轭的抗兔抗体。

蛋白质也可用放射性元素或酶来作标记。放射性标记可用任何现有的计量方法来检测。较佳的同位素可选自 ^3H 、 ^{14}C 、 ^{32}P 、 ^{35}S 、 ^{36}Cl 、 ^{51}Cr 、 ^{57}Co 、 ^{58}Co 、 ^{59}Fe 、 ^{90}Y 、 ^{125}I 、 ^{131}I 和 ^{186}Re 。

同样可使用酶标记, 其可用任何现在采用的比色法、分光法、荧光分光光度法、电流分析法或气体计量法技术来检测。酶通过与桥接分子如碳化二亚胺、二异氰酸酯、戊二醛等反应来共轭连接到选择的颗粒上。可用于这些步骤的许多酶是已知的并可被采用。较佳的是过氧化物酶、 β -葡萄糖醛酸酶、 β -D-葡萄糖苷酶、 β -D-半乳糖苷酶、脲酶、葡萄糖氧化酶加上过氧化物酶和碱性磷酸酶。美国专利 No. 3,654,090, 3,850,752 和 4,016,043 公开了不同标记物和方法的实施例。

该领域研究出并采用的特别测定系统是受体测定法。在受体测定法中, 对待测物作适当的标记, 然后对某种细胞测试菌落用一定量的标记接种, 然后进行结合研究以测定标记物与细胞受体的结合情况。通过这种方法就可确定物体间的亲和性差别。

该领域已知的有用的测定方法是“顺/反”测定法。简单的说, 该测定法采用了两种基因构建物, 一种通常是在转染入合适细胞系后可连续表达感兴趣的特定受体的质粒, 第二种是在受体/配体复合体的控制下表达报道分子如萤光素酶的质粒。因此, 例如, 如果要评价一种作为特定受体的配体的化合物, 一种质粒应是一种导致受体在所选细胞系中表达的构建物, 而第二种质粒应有与萤光素酶基因连接的启动子, 而萤光素酶基因中插入有特定受体的效应元件。如果待测化合物是受体的刺激物时, 配体将与受体复合, 得到的复合体与效应元件结合并启动萤

光素酶基因的转录。然后光计量测定得到的化学发光值，获得剂量效应曲线并与已知的配体比较。前述的过程在美国专利 No. 4,981,784 和 PCT 国际公开号 No. 88/03168 中有详细描述。

本发明公开了 STF-1 与 Pbx 结合，且该结合可增强促生长素抑制素和尽可能的胰岛素的转录。STF-1 与 Pbx 的结合需要五肽(FPWMK)。STF-1 和 Pbx 仅仅协同作用于能识别 STF-1-Pbx 异二聚体的亚型位点上。如下所示，STF-1 的氨基酸 115-123 需要有与 Pbx 的协同性。这就是说，由于突变的 STF-1 cDNA 不能与 Pbx 协同作用，因此需要 Pim。还显示了包括氨基酸残基 145-153 的 STF-1 同源异型域的柔性 N 端臂(art)对于与 Pbx 的协同性是必需的。

为描述本发明的各种例子给出了下列实施例，它们不意味着以任何方式来限制本发明。

实施例 1

质粒构建物和蛋白质的表达

图 3B 中显示的 STF-1 缺失的构建物以前已有描述(20)。STF-1 缺失的质粒用聚合酶链反应(PCR)扩增过程来构建。用 STF-1 cDNA 作为模板，使跨过缺失部分的有义寡核苷酸和与 STF-1 cDNA 3'端对应的反义寡核苷酸培育。用 Taq 聚合酶通过 PCR 扩增获得编码区内有合适缺失的 STF-1 cDNA 片段。用 Promega TNT 兔网织红细胞裂解物偶联转录/翻译系统根据生产商的方法制得 Pbx1 cDNA(C. Murre 所赠)和 STF-1 缺失突变型 D1-70、D1-115、D1-139 和 D216-284。在大肠杆菌中用细菌表达载体 pGEX3X 表达 STF-1 全长蛋白质和 STF-1 缺失突变体 hox140-215、1-140 和 110-215(见图 3A)。重组蛋白质如前所述(14)进行纯化；最初表达 STF-1 多肽作为融合蛋白，融合蛋白的 N 端含有谷胱甘肽-S-转移酶(GST)。GST-STF-1 多肽在谷胱甘肽琼脂糖凝胶树脂上用亲和层析从细菌裂解物中纯化获得。然后纯化后的蛋白质可用谷胱甘肽(10mM)从树脂上洗脱下来，或利用插入在 GST 和 STF-1 序列间的因子 X 断裂位点的蛋白裂解来获得。重组的蛋白质用 4μg 因子 Xa 在室温下消化 16 小时来从谷胱甘肽-琼脂糖凝胶粒上洗脱获得。

图 5 中描述的 GST-融合蛋白也用 pGEX3X 载体在大肠杆菌中表达。编码鼠 Isl1 同源异型域(12)(氨基酸 176 至 248)、仓鼠 cdx3 同源异型域(6)(氨基酸 143 至 253)和 STF-1 同源异型域(氨基酸 140 至 215)的序列用聚合酶链反应扩增并融合入 pGEX3X 质粒的框架中。对于 PBX 相互作用基序-同源异型域融合蛋白，STF-1

编码序列的残基 110-138(包括 FPWMK 基序)插入 pGEX3X 载体并与 Isl1 同源异型域(氨基酸 176 至 248)、cdx3 同源异型域(氨基酸 176 至 253)和 STF-1 同源异型域(氨基酸 140 至 215)的编码序列的上游框架连接。GST-融合蛋白如前所述(14)表达和纯化,并用 5mM 减少量的谷胱甘肽从珠粒上洗脱下来。

5

实施例 2

凝胶移位测定

凝胶移位测定用 0.1ng 的末端标记过的双链寡核苷酸(其序列表示在图 2C 中)如前所述(20)进行。对于解离速率(off-rate)分析, TSEII 探针首先在室温下与蛋白质一起培育 30 分钟,然后加入 1000 倍过量的未标记过的 TSEII 寡核苷酸,并将等份量在不同的时间点加样在凝胶上。在上移位测定(supershift assay)中,在加入 TSEII 探针前, TU6 核细胞抽提物在室温下预先用 1 μ l Pbx(11)或 STF-抗血清(20)培育 15 分钟。Pbx 抗血清不能区别 Pbx 家族内的成员(Pbx1,2,3)。

15

实施例 3

特异性蛋白质 DNA 复合体 C1、C2 和 C3

在胰肿瘤细胞系中,胰岛激素基因的促生长素抑制素和胰岛素表达与同源异型框因子 STF-1 有关。STF-1 通过与两种称为 TSEI 和 TSEII 的组织特异性调节元件结合来调节 Tu-6 细胞中促生长素抑制素的表达(14)。在 Tu-6 粗核抽提物的凝胶移位测定中,用 32 P-标记过的 TSEII 位点寡核苷酸探针可检测三种特异性蛋白质 DNA 复合体,它们在这里称为复合体 C1、复合体 C2 和复合体 C3(图 1A)。复合体 C1 和 C2 只发现存在于胰岛细胞抽提物中,而复合体 C3 可在非胰岛细胞系如 Hela 中检测到有相当水平。通过用 Oct-1 抗血清的“上移位测定”,表明复合体 C3 含有遍在 Oct-1 蛋白。相反,通过测得复合体 C1 和 C2 对加入的 STF-1 抗血清敏感表明复合体 C1 和 C2 含有 STF-1 蛋白。

在凝胶移位测定中,复合体 C2 随抽提物浓度呈指数型增长,而复合体 C1 呈线性增长。重组的全长 STF-1 蛋白质产生一个迁移至复合体 C1 相同位置处的复合体(图 1A,泳道 6),这表明复合体 C1 只含有 STF-1 蛋白,而复合体 C2 是高亲和性的异聚 STF-1 复合体。关于这一点,解离速率研究表明,与异聚的 STF-1 复合体 C2 约 15 分钟的半衰期相比,复合体 C1 的半衰期小于 1 分钟(图 1B)。

为确定复合体 C2 不与 STF-1 的同源二聚体型对应,在凝胶移位测定中用 TSEII 探针测定增加浓度的重组 STF-1 蛋白(图 1C)。在高含量的 STF-1 蛋白质下

没有发现有迁移较慢的复合体, 这表明复合体 C2 一定含有能使 STF-1 与 DNA 的结合大大稳定的其它蛋白质组分。

为证明该组分是普遍表达的, 将 Jurkat(或 Hela)粗核抽提物加入含有重组 STF-1 蛋白质的反应中。在这些条件下, 很容易发现有复合体 C2 形成, 且与加入重组 STF-1 有关(图 1C)。重建抽提物中的复合体 C2 的解离速率与 Tu-6 抽提物中的相当(约 15 分钟), 这表明遍在因子使 STF-1 和 TSEII 位点的结合稳定(图 1D)。

果蝇同源异型框蛋白(称为外小齿)与其它同源异型域蛋白协同结合到靶启动子位点上(2, 25)。下面测定复合体 C2 中是否含有哺乳动物同源外小齿(称为 Pbx)(23)。当将 Pbx 抗血清加入 Tu-6 粗核抽提物中, 前者可特异性封闭复合体 C2 的形成, 但是该抗血清对只含 STF-1 的 C1 的形成没有影响(图 2A, 泳道 2)。而且, Pbx 和重组 STF-1 蛋白的共同培育在凝胶移位测定中导致异聚 C2 复合体的形成(图 2B)。该 Pbx-STF-1 复合体的稳定性经解离速率测定, 与 Tu-6 核抽提物中的内源 C2 复合体相当(图 2C)。

为测定 Pbx 能否与 STF-1 协同结合到 TSEII 外的其它 STF-1 靶位点上, 用重组 STF-1 和 Pbx 蛋白在促生长素抑制素 TSEI 或胰岛素 P 和 FLAT 元件上进行凝胶移位测定, 与 STF-1 结合的位点和胰岛细胞所需的位点在胰细胞中分别限制了鼠促生长素抑制素和胰岛素 I 基因的表达(图 2D 和图 2E)。与 TSEII 元件相反(图 2D, 泳道 1-4), 在 TSEI、P 或 FLAT 元件上没有发现有异聚 STF-1-Pbx 复合体。实际上, 当 TSEI、P 或 FLAT 元件用于胰细胞粗抽提物的凝胶移位测定时, 异聚 STF-1 复合体也是不能被检测出的(数据未显示)(20)。

为测定 STF-1 和 Pbx 能否以协同方式刺激转录, 用报道基因载体进行瞬时转录, 该载体含有两个拷贝的位于最小生长激素启动子上游的促生长素抑制素 TSEII 位点(图 2F)。为优化转染的 Pbx 表达质粒的调控作用, 采用了可表达与 Pbx 融合的 E2A 激活域的 E2A-Pbx 载体(15)。当 STF-1 和 E2a-Pbx 分别转染入 GC 细胞中时, 它们对 TSE II 萤光素酶报道分子活性的影响可以忽略。然而, STF-1 和 E2a-Pbx 效应物质粒的共转染可显著提高 TSE II 报道分子的活性, 这表明 STF-1 实际上是与 Pbx 协同作用于 TSE II 位点上。相反, 用含有胰岛素 P 元件(P-Luc)的报道质粒没有发现有这种 STF-1 和 E2A-Pbx 间的协同性。由于 P 元件不能在 DNA 结合测定中形成 STF-1-Pbx 异二聚体(图 2D), 因此这些结果证明 STF-1 和 Pbx 仅仅协同作用于可识别 STF-1-Pbx 异二聚体的位点亚型上。

为鉴别对 Pbx-STF-1 协同性重要的促生长素抑制素 TSEII 元件的序列, 构建

了几个突变 TSEII 寡核苷酸(图 2D 和图 2E)。TSEII 位点含有 3 个 TAAT 基序, 而其是同源异型域蛋白的主要识别基序。TAAT 基序 1 或 3 的突变型(M1 和 M3, 正链)对 STF-1 和 Pbx 间协同结合的影响最小, 但 TAAT 基序 2(负链)完全清除了 Pbx-STF-1 的协同性, 这表明 STF-1-Pbx 复合体构成中那些特殊的 TAAT 基序是重要的。第三个 TAAT 基序的突变型(M3)也影响 STF-1 单聚体 C1 复合体的形成。

为鉴别 STF-1 中促进与 Pbx 协同结合的残基, 在凝胶移位测定中用 TSEII 探针测定一系列截短的 STF-1 多肽(图 3A 和图 3B)。C 端至同源异型框区域的残基缺失对与 Pbx 的协同结合性没有影响(D216-284 突变型)。N 端截短的缺少 STF-1 转录激活区(trans-activation domain)的 STF-1 多肽(氨基酸 1-115)也保留了结合到 TSEII 位点上的协同性。但再缺失氨基酸 115 至 123 的残基将使其丧失 Pbx 协同性, 这表明 STF-1 同源异型域外的区域对于异二聚体的形成是重要的(图 3A 和图 3B)。实际上, STF-1 的同源异型框区域(Hox140-215)在 TSEII 位点形成了单体复合体, 但不与 Pbx 协同结合。除 STF-1 同源异型域外还含有 N 端序列的 STF-1 多肽(氨基酸 110-215)表现出有 Pbx 协同性, 这表明与 STF-1 同源异型框侧接的 N 端残基可在与 TSEII 位点结合时与 Pbx 形成蛋白质-蛋白质接触。

在比较 STF-1 和果蝇的其它与外小齿(Pbx 果蝇同源物)协同结合 DNA 的同源异型域蛋白的过程中, 注意到 STF-1 的该 N 端区(氨基酸 115-123)有五肽基序 FPWMK, 其在各种后生动物的许多同源异型蛋白质中是保守的(5)。为描述该肽基序对于 Pbx 协同性的重要性, 构建一个突变型 STF-1 cDNA, 其基序中的每个残基均被氨基酸取代(FPWMK 变为 AAGGQ)(图 3C)。在凝胶移位测定中与野生型 STF-1 蛋白质比较, 诱变的 STF-1 蛋白质没有与重组 Pbx 或来自 Jurkat 抽提物的内源 Pbx 协同结合的能力(分别在图 3D 和图 3E 中)。相反, 突变型 STF-1 蛋白质在胰岛素 P 和 FLAT 元件上(不形成异二聚体复合体的位点处)显示出了野生型结合活性。这些结果证明 STF-1 中与 Pbx 相互作用的基序(称为 PBX 相互作用基序区或 pim)实际上是 Pbx 协同性所必需的。

为测定 PBX 相互作用基序区对于 Pbx 协同性来说是否足够, 测定该基序在融合入异源的同源异型框蛋白中时赋予 Pbx 协同性的能力(图 5B)。在凝胶移位测定中, lim 同源异型域蛋白 Isl1 可高亲和性地识别促生长素抑制素 TSEII 位点, 但是不能与 Pbx 异源二聚体化。很明显, Isl1 同源异型框上游的 PBX 相互作用基序区不能促进 Pbx 协同性。尾状因子 cdx3 含有与其同源异型域上游的基序 CEWMR 对应的 PBX 相互作用基序。但是, 同 Isl1 一样, cdx3 不与 Pbx 异源二聚体化; 且在 cdx3 同源异型域中加入 PBX 相互作用基序区不会促进 Pbx 协同性。

很明显, PBX 相互作用基序区和同源异型框之间的间距与 STF-1 和 PBX 相互作用基序-Isl1 以及 PBX 相互作用基序-cdx3 融合蛋白之间的间距是相当的, 这表明从同源异型框区域的间距不能解释它们的 Pbx 协同性丧失的原因。另外, 嵌合蛋白(PBX 相互作用基序-Isl1, PBX 相互作用基序-cdx3)可以以野生型亲和力与 DNA 结合, 这表明缺少协同性将不反映错误的折叠。综合考虑, 这些结果表明, 5 尽管 PBX 相互作用基序区是与 Pbx 形成异二聚复合体所需的, 但是 STF-1 同源异型域中的其它残基是协同结合到促生长素抑制素 TSE II 位点上所需的。

为表征 STF-1 同源异型域中除 PBX 相互作用基序区外的残基对于 Pbx 协同结合性在功能上是重要的, 将 cdx3 同源异型域中的片段用 STF-1 的相应片段代替(图 6C)。用 PBX 相互作用基序-cdx3 作为该实验的融合模板, 测得 STF-1 的 N 10 端臂减少了 Pbx 协同性, 但是 STF-1 同源异型域中的其它区域(helices 1,2 或 3)没有显示出这种活性。这些结果证明了 PBX 相互作用基序和 STF-1 的 N 端臂是与 Pbx 形成复合物所必需的。

Pbx 的果蝇同源物外小齿对于生长时靶基因如无翅(wingless)、teashirt、decapentaplegic 的亚型的合适激活是必需的(24)。外小齿通过与不同的同源异型蛋白质协同结合到 DNA 上来诱导这些靶基因(2, 25)。本发明证明外小齿的人同源物 Pbx 在促生长素抑制素 TSEII 元件上与 STF-1 形成了异二聚复合体。如凝胶移位测定中所测得的该复合体的解离速率降低, STF-1-Pbx 与单用 STF-1 比较是高度稳定的。然而, 重要的是应注意, 这些结果没有区别均明显有与 HOX 蛋白质协同结合的 Pbx 家族的不同成员(Pbx-1,2,3)。20

促生长素抑制素 TSEII 位点上 STF-1-Pbx 异二聚体的形成需要五肽基序(FPWMK), 而该基序在许多同源异型蛋白质中(不仅仅是脊椎动物, 而且在果蝇(5)和 C.elegans)是保守的。因此, 含有该基序的 C.elegans 同源异型蛋白质 mab5 或 lin-39 可同样与蛋白质 ceh-20(与 Pbx 有同源性)协同(1)。通过用人工靶序列来诱导 Pbx:hox 异二聚体形成, Chang 等人注意到该保守基序对于与 Pbx 协同结合 DNA 的重要性(3)。这些结果表明, Pbx 的作用不局限于 hox 复合体中的蛋白质, 25 也包括孤独同源异型框蛋白如 STF-1。

尽管 Chang 等人发现保守的 YPMWK 基序与 Hox A10 的融合足以促进与 Pbx 的协同性(3), 但是当该基序转移到胰岛细胞同源异型域蛋白如 isl-1 和 cdx-3 上 30 时, 没有发现诱导出协同性。这些结果表明同源异型域本身的其它残基是 STF-1/Pbx 复合体形成所需的。关于这一点, 发现 STF-1 同源异型域的柔性 N 端臂(氨基酸 145-153)对于 Pbx 协同性是必需的。N 端臂表现出赋予同源异型域蛋白如

Antennapedia 功能同源性, 尽管潜在机理仍是未知的(4; 26), 结构研究揭示了 N 端臂位于 DNA 的小沟处, 在那里它提供了 DNA 结合或蛋白质-蛋白质相互作用时的微小区别。尽管这些结果没有对这些模型作出区别, 但是, 它试图推测 STF-1/Pbx 复合体的形成与 N 端臂形成同源性小沟接触部位的能力是部分有关的。

STF-1-Pbx 异二聚体的形成仅仅发生在潜在 STF-1 靶位点的亚型上。这些结果表明, 这一偏爱可能形成了胰岛细胞生长中靶部位的选择性基础。以前就注意到, 例如, 在胰岛的不同细胞(分别为 β 和 d)中, STF-1 诱导胰岛素和促生长素抑制素两者的表达。在 β 细胞中, STF-1 可通过与螺旋环螺旋蛋白质 E47 协同作用诱导胰岛素表达。相反, STF-1 通过与 Pbx 协同结合到 TSEII 位点上来促进 d 细胞中的表达。这些观察结果表明, 帮助胰岛谱系中的细胞表达胰岛素或促生长素抑制素都与 E 框与 Pbx 型蛋白结合的相对表达有关。

和这里对胰生长的预想一样, 也描述了酵母中生长调控的组合机理。关于这一点, 同源异型域蛋白 Mata2 与二倍体/细胞中的 Mata1 协同结合 hsg 操纵基因并抑制单倍体特异性基因。但是在单倍体或细胞中, Mata2 表现出与不同的激活物 MCM1 协同, 从而来激活不同的基因程序(9)。

使得 Pbx 和某些同源异型框蛋白质间能相互作用的保守基序的存在可部分解释生长中该调控器的全局性作用。除了胰外, 促生长素抑制素还在许多组织中如脑、胃和髓甲状腺中表达。尽管在这些组织中能控制促生长素抑制素表达的因子还未被证实, 但是该结果表明这些蛋白质可通过与 Pbx 形成异二聚复合体来刺激 TSEII 元件上促生长素抑制素的表达。

这里引用了下列文献:

1. Burglin T.R. 等人, *Nature Genet.* 1:319-320, 1992.

2. Chan S. 等人, *Cell* 78:603-615, 1994.

3. Chang C. 等人, *Cenes & Dev.* 9:663-674, 1995.

4. Furukubo-Tokunaga K. 等人, *Proc. Natl. Acad. Sci.* 90:6360-6364, 1994.

5. Gehring W.J. 等人, *Annu. Rev. Biochem.* 63:437-526, 1994.

6. German M.S. 等人, *Genes & Dev.* 6:2165-2176, 1992.

7. Guz Y. 等人, *Development* 121:11-18, 1995.

8. Hayashi S. 等人, *Cell* 63:883-894, 1990.

9. Johnson A., Acombinatorial regulatory circuit in budding yeast., In S. McKnight(ed.), *transcriptional regulation*, Cold Spring Harbor Lab. Press, NY, p.

975-1006, 1992.

10. Jonsson J. 等人, *Nature* 371:606-609, 1994.
11. Kamps M. P. 等人, *Genes & Dev.* 5:358-368, 1991.
12. Karlsson O. 等人, *Nature* 344:879-882, 1990.
- 5 13. Krumlauf R., *Cell* 78:191-201, 1994.
14. Leonard J. 等人, *Mol. Endocrinol.* 7:1275-1283, 1993.
15. Lu Q. 等人, *Mol. Cell. Biol.* 14:3938-3948, 1994.
16. McGinnis W. 等人, *Cell* 68:283-302, 1992.
17. Miller C.P. 等人, *EMBO J.* 13:1145-1156, 1994.
- 10 18. Moncia K. 等人, *Mol. Cell. Biol.* 11:6149-6157, 1991.
19. Ohlsson H. 等人, *EMBO J.* 12:4251-4259, 1993.
20. Peers B. 等人, *Mol. Endocrinol.* 8:1798-1806, 1994.
21. Peifer M. 等人, *Genes & Dev.* 4:1209-1223, 1990.
22. Peshavaria M. 等人, *Mol. Endocrinol.* 8:806-816, 1994.
- 15 23. Rauskolb C., 等人, *Cell* 74:1101-1112, 1993.
24. Rauskolb C., 等人, *EMBO J.* 13:3561-3569, 1994.
25. van Dijk M., 等人, *Cell* 78:616-624, 1994.
26. zeng W., 等人, *Development* 118:339-352, 1993.

20 本说明书中提及的任何专利或出版物表现出本发明所属领域技术人员的水
平。这些专利和出版物在这里以相同程度作引证参考，就如每个单独的出版物特
别单独结合入本发明作参考一样。

该领域技术人员应容易理解本发明能够实现所提到的以及潜在的目的、最终
产品和优点。本发明的实施例以及这里描述的方法、步骤、处理、分子和特定的
化合物是较佳实施例的代表，是典型的而不限本发明的范围。该领域技术人员
25 可在有权利要求范围确定的本发明精神包括范围内作变化和其它应用。

序列表

- (1) 一般信息:
- 5 (i) 申请人: Montminy, Marc 和 Peers, Bernard
- (ii) 发明名称: 刺激促生长素抑制素和胰岛素产生的化合物的筛选方法
- (iii) 序列数目: 15
- (iv) 通信地址:
- (A) 收信人: Sarah Brashears, MCGRGOR & ADLER, P.C.
- 10 (B) 街道: 8011 Candle Lane
- (C) 城市: Houston
- (D) 州: Texas
- (E) 国家: USA
- (F) 邮编: 77071
- 15 (v) 计算机可读形式:
- (A) 记录介质类型: 软盘
- (B) 计算机: Apple Macintosh
- (C) 操作系统: Macintosh
- (D) 软件: Microsoft Word for Macintosh
- 20 (vi) 本申请资料:
- (A) 申请号:
- (B) 申请日:
- (C) 分类:
- (viii) 律师/代理人信息:
- 25 (A) 姓名: Brashears, Sarah J.
- (B) 登记号: 38,087
- (C) 参考/案卷号: D-5849
- (ix) 通讯信息:
- (A) 电话: 713-777-2321
- 30 (B) 传真: 713-777-6908

(2) SEQ ID NO:1 的信息:

- 5 (i) 序列特征:
(A) 长度: 22bp
(B) 类型: 核酸
(C) 链型: 双链
(D) 拓扑结构: 线性
- (ii) 分子类型:
(A) 描述: 其它核酸
- (iii) 假设: 否
- (iv) 反义: 否
- 10 (vi) 最初来源:
(B) 菌株:
(C) 单独分离物:
(D) 发展阶段:
(F) 组织型:
- 15 (G) 细胞型:
(H) 细胞系:
- (xi) 序列描述: SEQ ID NO:1:
GATCTCAGTA ATTAATCATG CA 22
- 20 (3) SEQ ID NO:2 的信息:
- (i) 序列特征:
(A) 长度: 22bp
(B) 类型: 核酸
(C) 链型: 双链
- 25 (D) 拓扑结构: 线性
- (ii) 分子类型:
(A) 描述: 其它核酸
- (iii) 假设: 否
- (iv) 反义: 否
- 30 (vi) 最初来源:
(B) 菌株:
(C) 单独分离物:

(D) 发展阶段:

(F) 组织型:

(G) 细胞型:

(H) 细胞系:

5 (xi) 序列描述: SEQ ID NO:2:
AGTCATTAAT TAGTACGTCA TG 22

(4) SEQ ID NO:3 的信息:

(i) 序列特征:

10 (A) 长度: 22
(B) 类型: 核酸
(C) 链型: 双链
(D) 拓扑结构: 线性

(ii) 分子类型:

15 (A) 描述: 其它核酸

(iii) 假设: 否

(iv) 反义: 否

(vi) 最初来源:

(B) 菌株:

20 (C) 单独分离物:

(D) 发展阶段:

(F) 组织型:

(G) 细胞型:

(H) 细胞系:

25 (xi) 序列描述: SEQ ID NO:3:
GATCTCAGC TATTAATCATG CA 22

(5) SEQ ID NO:4 的信息:

(i) 序列特征:

30 (A) 长度: 22bp
(B) 类型: 核酸
(C) 链型: 双链

- (D) 拓扑结构: 线性
- (ii) 分子类型:
- (A) 描述: 其它核酸
- (iii) 假设: 否
- 5 (iv) 反义: 否
- (vi) 最初来源:
- (B) 菌株:
- (C) 单独分离物:
- (D) 发展阶段:
- 10 (F) 组织型:
- (G) 细胞型:
- (H) 细胞系:
- (xi) 序列描述: SEQ ID NO:4:
- AGTCGATAAT TAGTACCTCA TC 22
- 15
- (6) SEQ ID NO:5 的信息:
- (i) 序列特征:
- (A) 长度: 23
- (B) 类型: 核酸
- 20 (C) 链型: 双链
- (D) 拓扑结构: 线性
- (ii) 分子类型:
- (A) 描述: 其它核酸
- (iii) 假设: 否
- 25 (iv) 反义: 否
- (vi) 最初来源:
- (B) 菌株:
- (C) 单独分离物:
- (D) 发展阶段:
- 30 (F) 组织型:
- (G) 细胞型:
- (H) 细胞系:

(xi) 序列描述: SEQ ID NO:5:
GATCTCAGTAA TCTAATCATG CA 23

(7) SEQ ID NO:6 的信息:

- 5 (i) 序列特征:
(A) 长度: 22
(B) 类型: 核酸
(C) 链型: 双链
(D) 拓扑结构: 线性
- 10 (ii) 分子类型:
(A) 描述: 其它核酸
(iii) 假设: 否
(iv) 反义: 否
(vi) 最初来源:
- 15 (B) 菌株:
(C) 单独分离物:
(D) 发展阶段:
(F) 组织型:
(G) 细胞型:
- 20 (H) 细胞系:

(xi) 序列描述: SEQ ID NO:6:
AGTCATTAAT ATGTACGTCA TG 22

(8) SEQ ID NO:7 的信息:

- 25 (i) 序列特征:
(A) 长度: 22bp
(B) 类型: 核酸
(C) 链型: 双链
(D) 拓扑结构: 线性
- 30 (ii) 分子类型:
(A) 描述: 其它核酸
(iii) 假设: 否

(iv) 反义: 否
(vi) 最初来源:
(B) 菌株:
(C) 单独分离物:
5 (D) 发展阶段:
(F) 组织型:
(G) 细胞型:
(H) 细胞系:
(xi) 序列描述: SEQ ID NO:7:
10 GATCTCAGTA ATTATACATG CA 22

(9) SEQ ID NO:8 的信息:
(i) 序列特征:
(A) 长度: 22bp
15 (B) 类型: 核酸
(C) 链型: 双链
(D) 拓扑结构: 线性
(ii) 分子类型:
(A) 描述: 其它核酸
20 (iii) 假设: 否
(iv) 反义: 否
(vi) 最初来源:
(B) 菌株:
(C) 单独分离物:
25 (D) 发展阶段:
(F) 组织型:
(G) 细胞型:
(H) 细胞系:
(xi) 序列描述: SEQ ID NO:8:
30 AGTCATTAAT ATGTACGTCA TG 22

(10) SEQ ID NO:9 的信息:

- 5 (i) 序列特征:
(A) 长度: 25bp
(B) 类型: 核酸
(C) 链型: 双链
(D) 拓扑结构: 线性
- 10 (ii) 分子类型:
(A) 描述: 其它核酸
(iii) 假设: 否
(iv) 反义: 否
(vi) 最初来源:
(B) 菌株:
(C) 单独分离物:
(D) 发展阶段:
(F) 组织型:
(G) 细胞型:
(H) 细胞系:
- 15 (xi) 序列描述: SEQ ID NO:9:
GATCTAGAGC CCTTAATGGG CCAAA 25
- 20 (11) SEQ ID NO:10 的信息:
(i) 序列特征:
(A) 长度: 25bp
(B) 类型: 核酸
(C) 链型: 双链
(D) 拓扑结构: 线性
- 25 (ii) 分子类型:
(A) 描述: 其它核酸
(iii) 假设: 否
(iv) 反义: 否
- 30 (vi) 最初来源:
(B) 菌株:
(C) 单独分离物:

(D) 发展阶段:

(F) 组织型:

(G) 细胞型:

(H) 细胞系:

5 (xi) 序列描述: SEQ ID NO:10:
ATCTCGGGAA TTACCCGGTT TCTAG 25

(12) SEQ ID NO:11 的信息:

(i) 序列特征:

10 (A) 长度: 23bp
(B) 类型: 核酸
(C) 链型: 双链
(D) 拓扑结构: 线性

(ii) 分子类型:

15 (A) 描述: 其它核酸

(iii) 假设: 否

(iv) 反义: 否

(vi) 最初来源:

(B) 菌株:

20 (C) 单独分离物:

(D) 发展阶段:

(F) 组织型:

(G) 细胞型:

(H) 细胞系:

25 (xi) 序列描述: SEQ ID NO:11:
GATCTTGCGA GGCTAATGGT GCG 23

(13) SEQ ID NO:12 的信息:

(i) 序列特征:

30 (A) 长度: 23bp
(B) 类型: 核酸
(C) 链型: 双链

- (D) 拓扑结构: 线性
- (ii) 分子类型:
- (A) 描述: 其它核酸
- (iii) 假设: 否
- 5 (iv) 反义: 否
- (vi) 最初来源:
- (B) 菌株:
- (C) 单独分离物:
- (D) 发展阶段:
- 10 (F) 组织型:
- (G) 细胞型:
- (H) 细胞系:
- (xi) 序列描述: SEQ ID NO:12:
- AACGCTCCGA TTACCACGCC ATG 23
- 15
- (14) SEQ ID NO:13 的信息:
- (i) 序列特征:
- (A) 长度: 28bp
- (B) 类型: 核酸
- 20 (C) 链型: 双链
- (D) 拓扑结构: 线性
- (ii) 分子类型:
- (A) 描述: 其它核酸
- (iii) 假设: 否
- 25 (iv) 反义: 否
- (vi) 最初来源:
- (B) 菌株:
- (C) 单独分离物:
- (D) 发展阶段:
- 30 (F) 组织型:
- (G) 细胞型:
- (H) 细胞系:

(xi) 序列描述: SEQ ID NO:13:
GATCTTGTTA ATAATCTAAT TACCCTAG 28

(15) SEQ ID NO:14 的信息:

- 5 (i) 序列特征:
(A) 长度: 27bp
(B) 类型: 核酸
(C) 链型: 双链
(D) 拓扑结构: 线性
- 10 (ii) 分子类型:
(A) 描述: 其它核酸
(iii) 假设: 否
(iv) 反义: 否
(vi) 最初来源:
- 15 (B) 菌株:
(C) 单独分离物:
(D) 发展阶段:
(F) 组织型:
(G) 细胞型:
- 20 (H) 细胞系:

(xi) 序列描述: SEQ ID NO:14:
AACAAATTATT AGATTAATGG GATCCAT 27

(16) SEQ ID NO:15 的信息:

- 25 (i) 序列特征:
(A) 长度: 5 碱基
(B) 类型: 氨基酸
(C) 链型单链
(D) 拓扑结构: 线性
- 30 (ii) 分子类型:
(A) 描述: 肽
(iii) 假设: 否

- 5
- (iv) 反义: 否

(vi) 最初来源:

(B) 菌株:

(C) 单独分离物:

(D) 发展阶段:

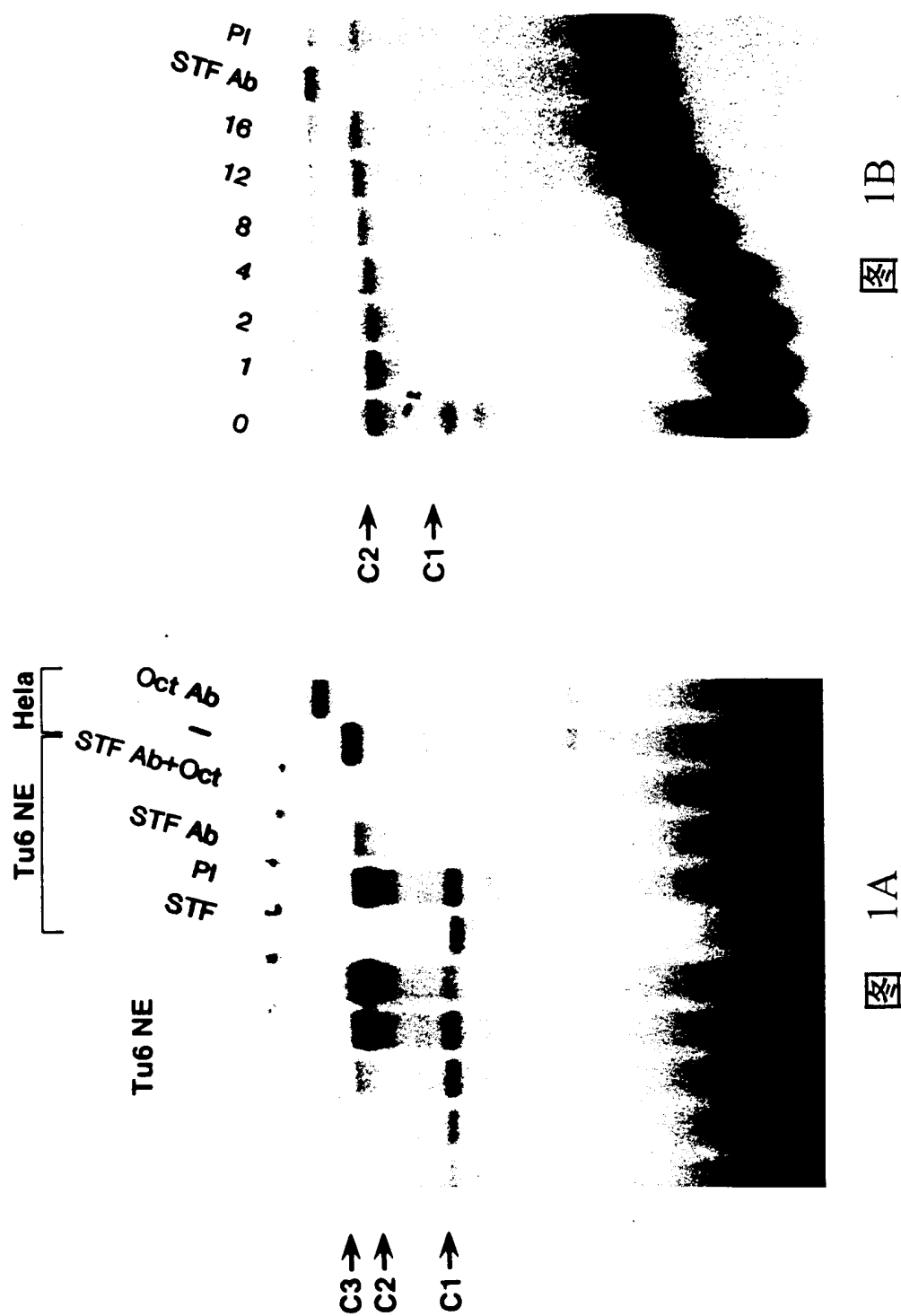
(F) 组织型:

(G) 细胞型:

(H) 细胞系:

(xi) 序列描述: SEQ ID NO:15:
- 10
- Phe Pro Trp Met Lys

5



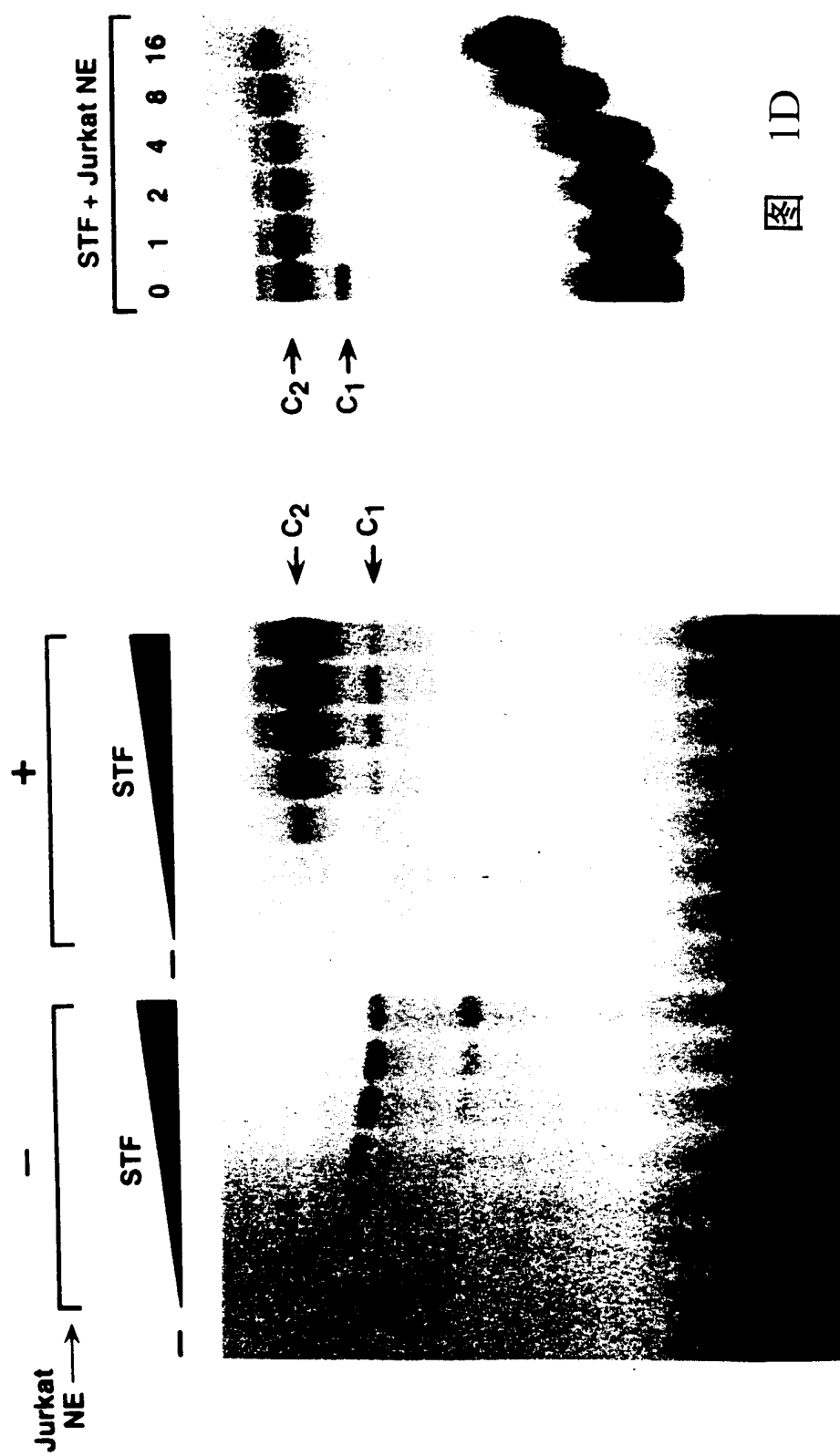


图 1C

图 1D

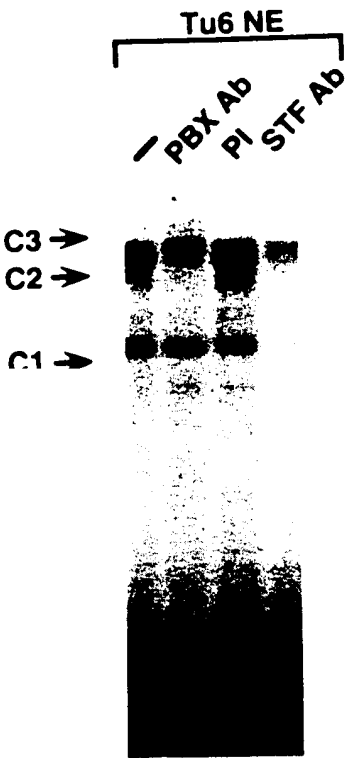


图 2A

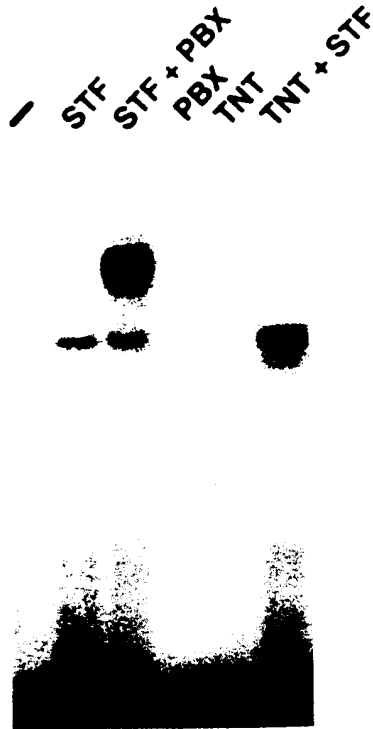


图 2B

0 1 2 4 6 8 12 16

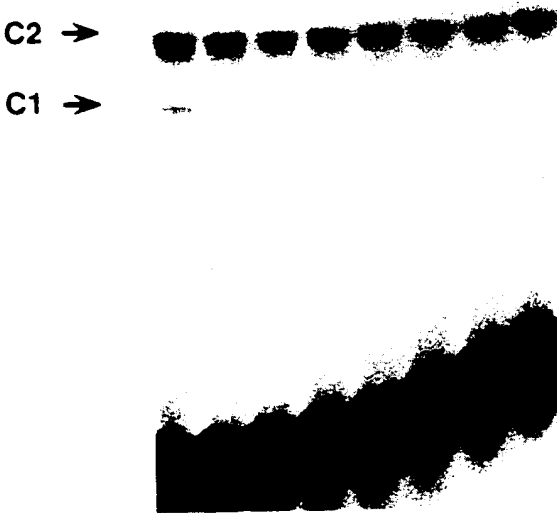


图 2C

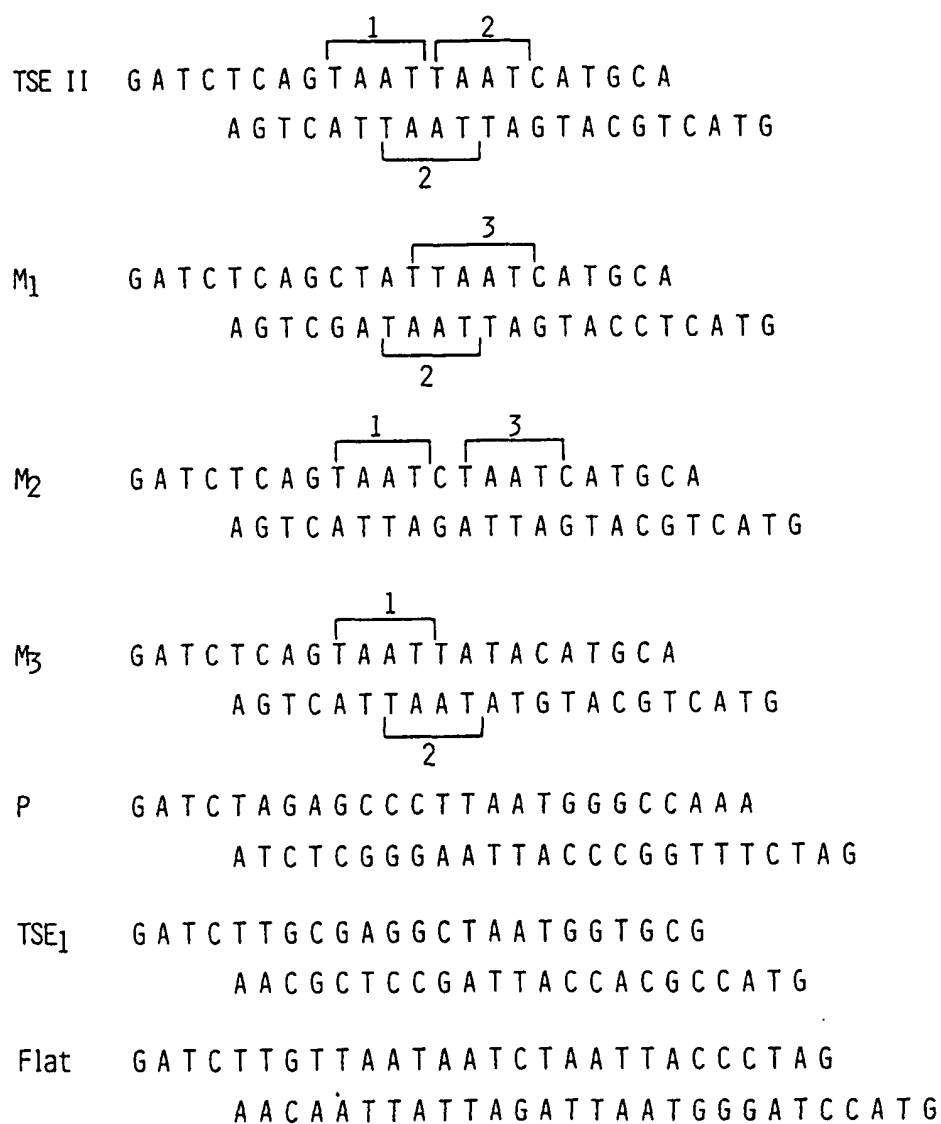


图 2E

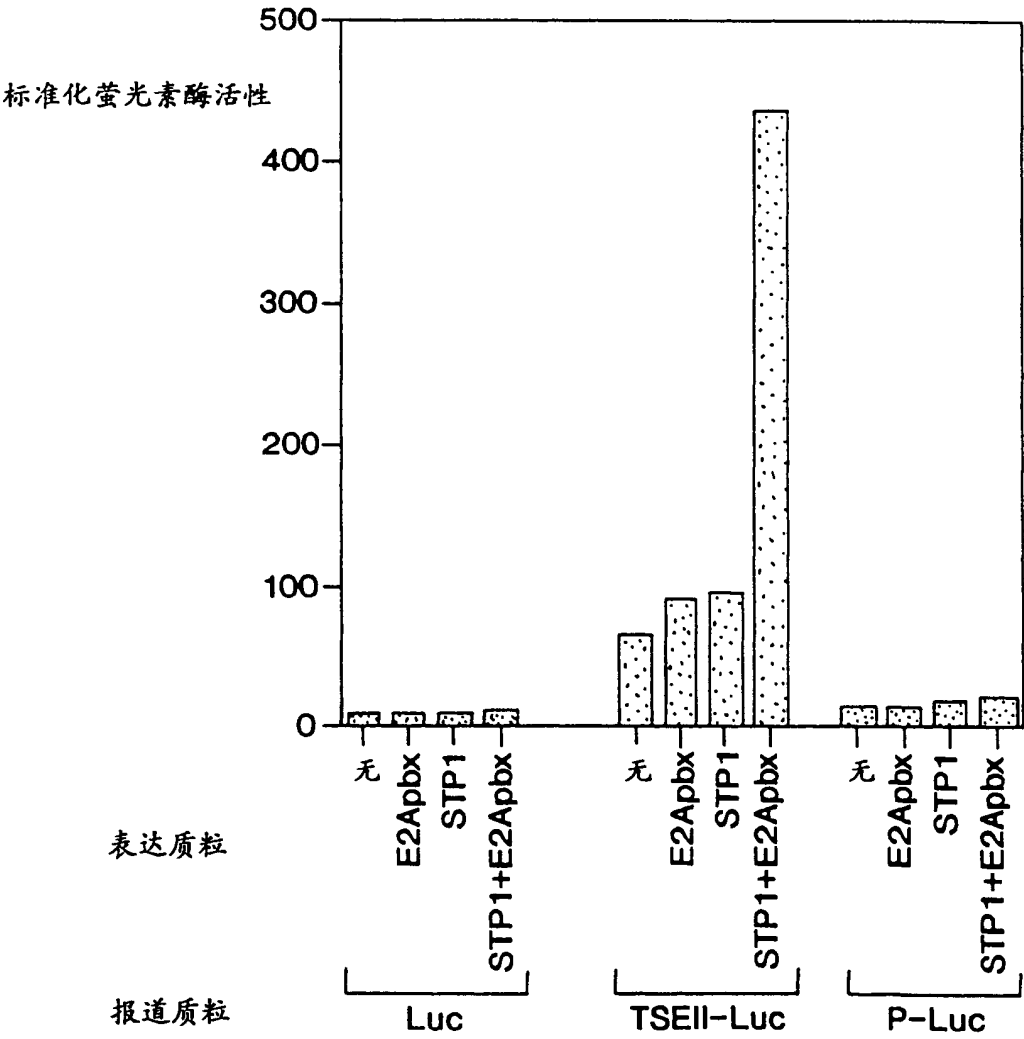


图 2F

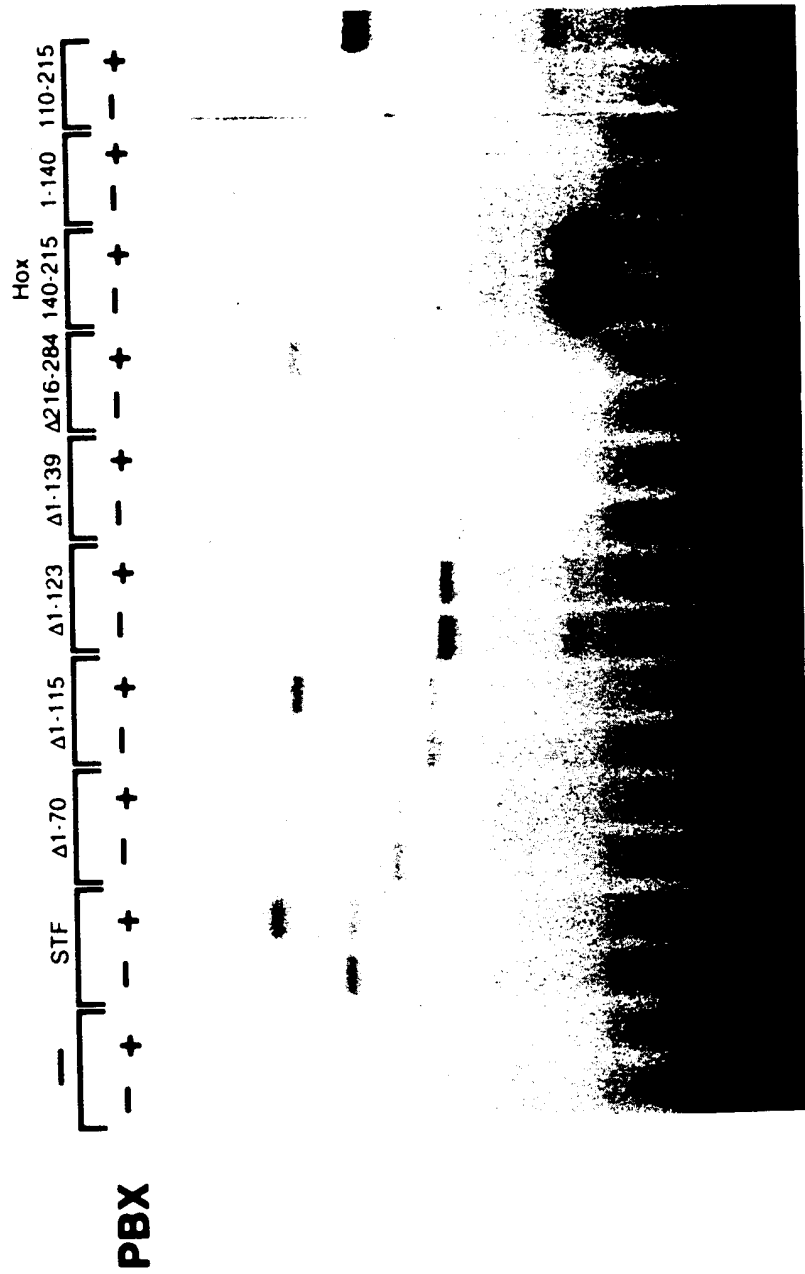


图 3A

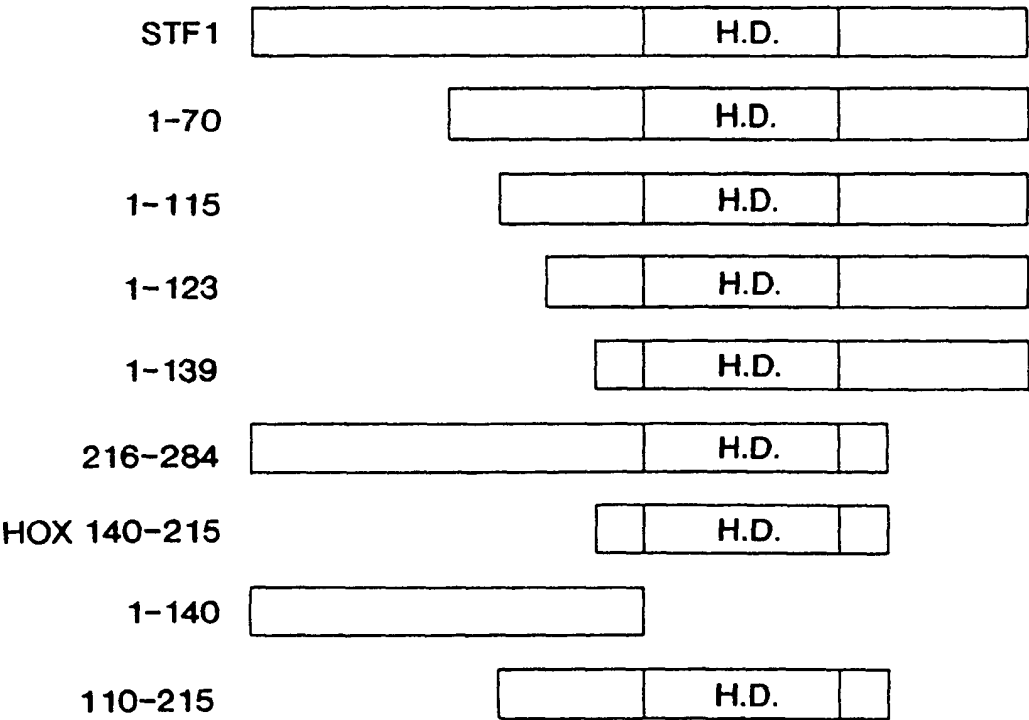


图 3B

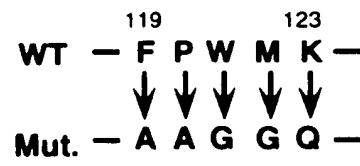


图 3C

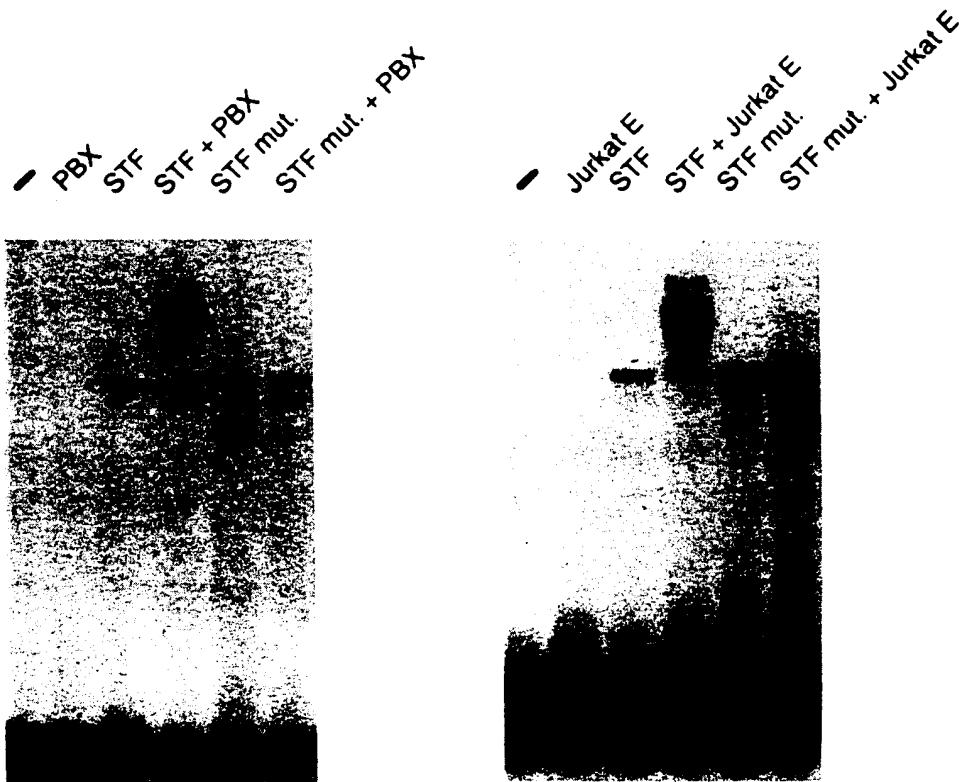


图 3D

图 3E

	名称	基序	同源异型域至 N端距离(残基 个数)
鼠	STF ₁	FPWMK	22
果蝇 (ANT和BX复合体)	Labial (lab) Proboscipedia (pb) Deformed (dfd) Sex Combs reduced (Scr) Antennapedia (Antp) Ultrabithorax (Ubx) Abdominal-A (AbdA)	YKWMQ YPWMK YPWMK YPWMK YPWMR YPMMA YPMMT	120 28 17 14 8 50 24
果蝇 (孤独同源异型框)	Caudal (cad)	FDWMK	15
Celegans (HOM复合体)	mab 5 lin 39	FPWMK YPMMT	8 11
人 (Hox复合体)	Hox B1 Hox B2 Hox B3 Hox B4 Hox B5 Hox B6 Hox B7 Hox B8	FDWMK FPWMK FPWMK YPWMR FPWMR YPWMQ YPWMR FPWMR	18 43 53 15 12 13 5 6
哺乳动物 (孤独同源异型框)	Hox 11 m cdx ₁ m cdx ₂ m cdx ₄	FPWMK YAWMR CEWMR YAWMR	23 11 15
Sea urchin	Hox ₁	YPWMK	11
	共有序列	Y <u>PWM</u> K F <u>PWM</u> R	

图 4

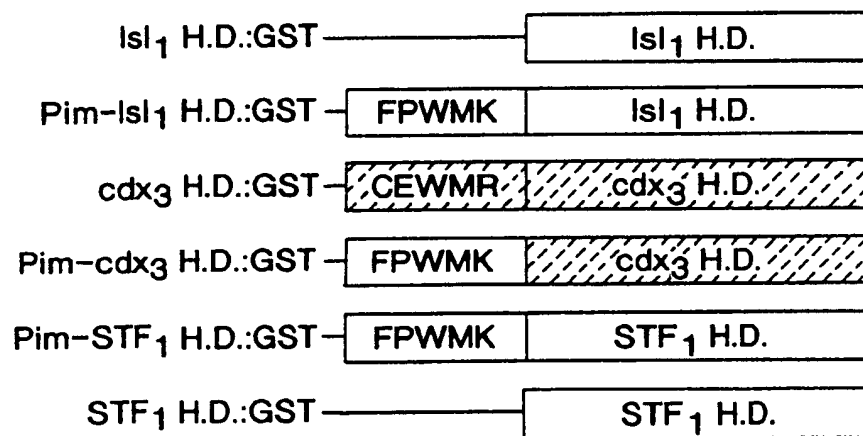


图 5A

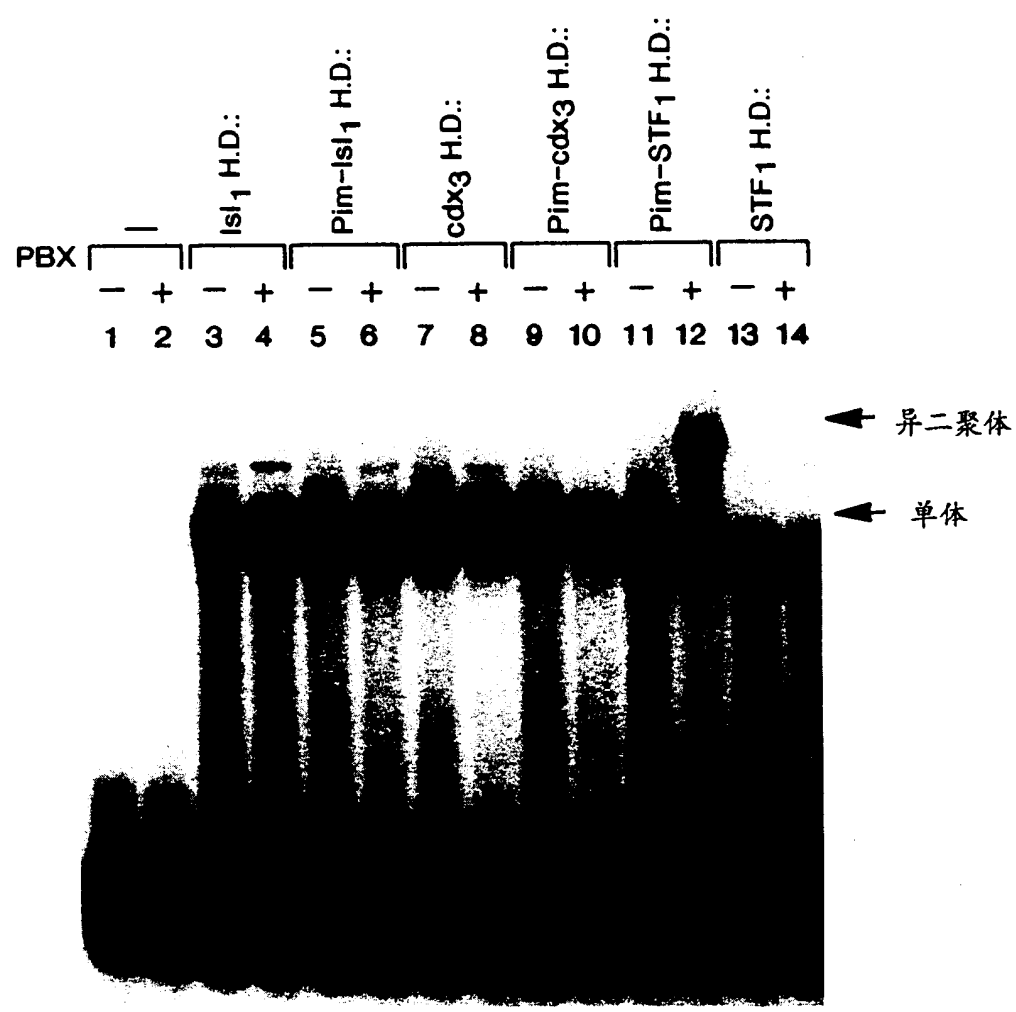


图 5B

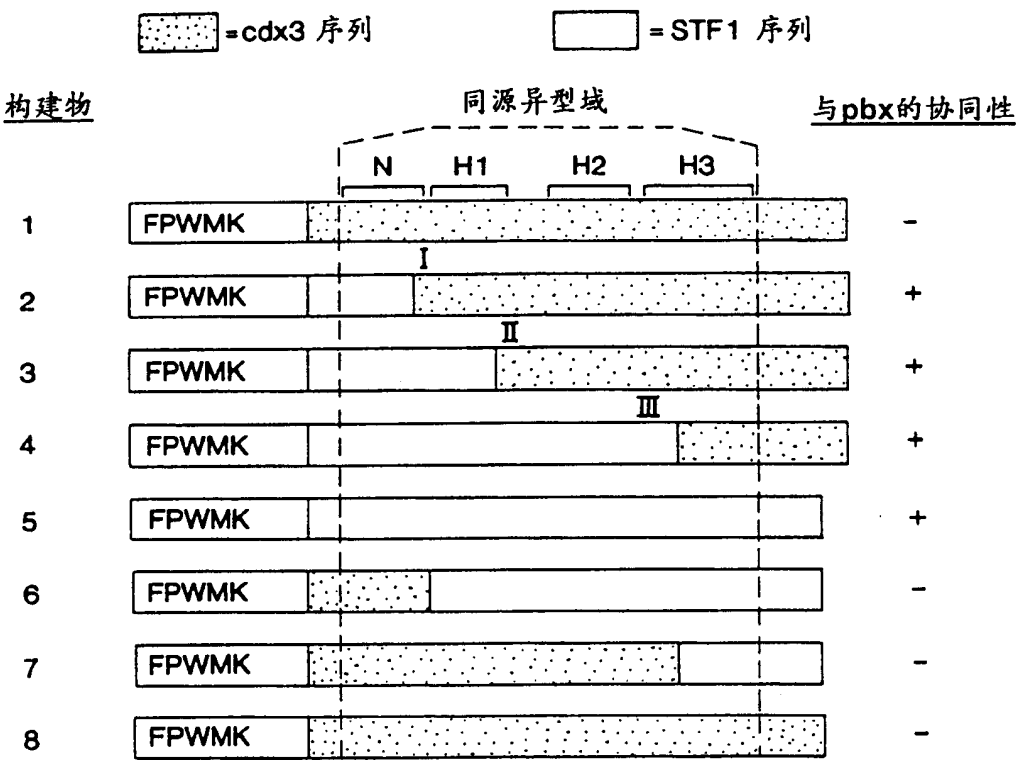


图 6A

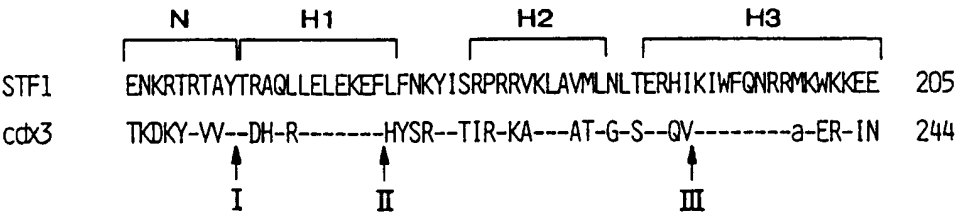


图 6B

构建物	—		1	2	3	4	5	6	7	8
PBX		+	+		+		+	+	+	+
TNT		+	+	+		+	+			

C2 ↑
C1 ↑



图 6C